Semi-Discrete approximation of Optimal Mass Transport

G. Wolansky,
Department of Mathematics, Technion, Haifa 32000, Israel [1]

Abstract

Optimal mass transport is described by an approximation of transport cost via semi-discrete costs. The notions of optimal partition and optimal strong partition are given as well. We also suggest an algorithm for computation of Optimal Transport for general cost functions induced by an action, an asymptotic error estimate and several numerical examples of optimal partitions.

1 Introduction

Optimal mass transport (OMT) goes back to the pioneering paper of Monge [15] at the 18th century. In 1942, L. Kantorovich [13] observed that OMT can be relaxed into an infinite dimensional linear programming in measure spaces. As such , it has a dual formulation which is very powerful and was later (1987) used by Brenier [3] to develop the theory of Polar factorization of positive measures. OMT has many connections with PDE, kinetic theory, fluid dynamics, geometric inequalities, probability and many other fields in mathematics as well as in computer science and economy.

Even though finite dimensional (or discrete) OMT is well understood, its extension to infinite dimensional measure spaces poses a great challenge, e.g. uniqueness and regularity theory of fully non-linear PDE such as the Monge-Amper equation [6].

We suggest to investigate a bridge between finite (”discrete”) and infinite (”continuum”) dimensional OMT. This notion of *semi-discrete* OMT leads naturally to *optimal partition* of measure spaces. Our motivation in this paper is the development of numerical method for solving OMT. Efficient algorithms are of great interest to many fields in operational research and, recently, also for optical design [9, 19, 20] and computer vision (”earth moving metric”) [21].

When dealing with numerical approximations for OMT, the problem must be reduced to a discrete, finite OMT (with, perhaps, very large number of degrees of freedom). Discrete OMT is often called the *assignment problem.* This is, in fact, a general title for a variety of linear and quadratic programming. It seems that the first efficient algorithm was the so called ”Hungarian Algorithm”, after two Hungarian mathematicians. See [11, 23, 12, 8, 16] and the survey paper [18] for many other relevant references.

The deterministic, finite assignment problem is easy to formulate. We are given n men and n women. The cost of matching man i to a woman j is $c_{i,j}$. The object is to find the assignment (matching) $i \to j$, given in terms of a permutation $j = \tau(i)$ which minimize the total cost of matching $\sum_{i=1}^n c_{i,\tau(i)}$.

When replacing the deterministic assignment by a probabilistic one, we assign the probability $p_i^j \geq 0$ for matching man i to woman j. The discrete assignment problem is then reduced to the

[1] Email: gershonw@math.tecnion.ac.il

linear programming of minimizing

$$\sum_{i=1}^{n}\sum_{j=1}^{n} p_i^j c_{i,j} \tag{1}$$

over all stochastic $n \times n$ matrices $P := \{p_i^j\}$, i.e. these matrices which satisfy the $2n + n^2$ linear constraints

$$\sum_{k=1}^{n} p_k^j = \sum_{k=1}^{n} p_i^k = 1 \ \ ; \ \ p_i^j \geq 0 \ \ \forall\, i, j \in \{1, \ldots n\} \ .$$

The Birkhoff Theorem assures us, to our advantage, that the optimal solution of this continuous assignment problem is also the solution of the deterministic version.

The probabilistic version seems to be more difficult since it involves a search on a much larger set of $n \times n$ stochastic matrices. On the other hand, it has a clear advantage since it is, in fact, a linear programming which can be handled effectively by well developed algorithms for such problems.

In many cases the probabilistic version cannot be reduced to the deterministic problem. For example, if the number of sources n and number of targets m not necessarily equal, or when not all sources must find target, and/or not all targets must be met, then the constraints are relaxed into $\sum_{i=1}^{n} p_i^j \leq 1$ and/or $\sum_{i=1}^{m} p_j^i \leq 1$. We shall not deal with these extension in the current paper, except, to some extent, in section 4 below.

1.1 From the discrete assignment problem to the continuum OMT

Let μ be a probability measure on some measure space X, and ν another probability measure on (possibly different) measure space Y. Let $c = c(x, y)$ be the cost of transporting x to y. The object of the Monge problem is to find a measurable mapping $T : X \to Y$ which generalizes the deterministic assignment perturbation τ described above in the following sense:

$$T_{\#}\mu = \nu \ \text{ namely } \ \mu(T^{-1}(B)) = \nu(B) \tag{2}$$

for every $\nu-$measurable set $B \subset Y$. The optimal Monge mapping (if exists) realizes the infimum

$$\inf_{T_{\#}\mu=\nu} \int_X c(x, T(x))\mu(dx) \ .$$

The relaxation of Monge problem into Kantorovich problem is analogues to the relaxation of the deterministic assignment problem to the probabilistic one: Find the minimizer

$$c(\mu, \nu) := \min_{\pi \in \Pi_X^Y(\mu,\nu)} \int_X \int_Y c(x, y)\pi(dxdy) \tag{3}$$

among all probability measures $\pi \in \Pi_X^Y(\mu, \nu) :=$

$$\{ \text{ Probability measures on } X \times Y \text{ whose } X \text{ (resp. } Y\text{) marginals are } \mu \text{ (resp. } \nu)\} \ . \tag{4}$$

In fact, Kantorovich problem is just an infinite dimensional linear programming over the huge set $\Pi_X^Y(\mu, \nu)$. The Monge problem can be viewed as a restriction of the Kantorovich problem to the class of *deterministic* probability measures in $\Pi_X^Y(\mu, \nu)$, given by $\pi(dxdy) = \mu(dx)\delta_{y-T(x)}$ where $T_{\#}\mu = \nu$. It turns out, somewhat surprisingly, that the value $c(\mu, \nu)$ of the Kantorovich problem equals to the infimum (3) of Monge problem, provided c is a continuous function on $X \times Y$ and μ does not contain a Dirac δ singularity (an atom) [1].

1.2 Semi-finite approximation- The middle way

Suppose the transportation cost $c = c(x, y)$ on $X \times Y$ can be obtained by interpolation of pair of functions $c^{(1)}$ on $X \times Z$ and $c^{(2)}$ on $Z \times Y$, where Z is a third domain and the interpolation means

$$c(x, y) := \inf_{z \in Z} c^{(1)}(x, z) + c^{(2)}(z, y) \ . \tag{5}$$

A canonical example for $X = Y = \mathbb{R}^d$ is $c(x, y) = c(|x - y|)$ where $c(w) = |w|^p$, $p \geq 1$. Then (5) is valid for $Z = \mathbb{R}^d$ and both $c^{(1,2)}(w) = 2^{p-1}|w|^p$. So

$$c(x, y) := |x - y|^p = 2^{p-1} \inf_{z \in \mathbb{R}^d} |x - z|^p + |z - y|^p \tag{6}$$

for any $x, y \in \mathbb{R}^d$ provided $p \geq 1$. Note in particular that the minimizer above is unique, $z = (x + y)/2$, provided $p > 1$, while $z = tx + (1 - t)y$ for any $t \in [0, 1]$ if $p = 1$.

Let $Z = Z_m := \{z_1, \ldots z_m\} \subset Z$ is a finite set. Denote

$$c^{Z_m}(x, y) := \min_{z \in Z_m} c^{(1)}(x, z) + c^{(2)}(z, y) \geq c(x, y) \tag{7}$$

the (Z_m) semi-finite approximation of c given by (5).

An optimal transport plan for a semi-discrete cost (7) is obtained as a pair of $m-$*partitions* of the spaces X and Y. An $m-$partition is a decomposition of the the space into m mesurable, mutually disjoint subset. It turns out that $c^{Z_m}(\mu, \nu)$ can be obtained as

$$c^{Z_m}(\mu, \nu) = \inf_{\{A_z\},\{B_z\}} \sum_{z \in Z_m} \int_{A_z} c^{(1)}(x, z)\mu(dx) + \int_{B_z} c^{(2)}(z, y)\nu(dz) \tag{8}$$

where the infimum is on the pair of partitions $\{A_z\}$ of X and $\{B_z\}$ of Y satisfying $\mu(A_z) = \nu(B_z)$ for any $z \in Z_m$. The optimal plan is, then, reduced to m plans transporting $\bar{A}_z \subset X$ to $\bar{B}_z \subset Y$, for any $z \in Z_m$, where $\{\bar{A}_z, \bar{B}_z\}$ is the optimal partition realizing (8).

The real advantage of the semi-discrete method described above is that it has a dual formulation which convert the optimization (8) to a convex optimization on $\mathbb{R}^m$. Indeed, we prove that for a given $Z_m \subset Z$ there exists a concave function $\Xi^\nu_{\mu,Z_m} : \mathbb{R}^m \to \mathbb{R}$ such that

$$\max_{\vec{p} \in \mathbb{R}^m} \Xi^\nu_{\mu,Z_m}(\vec{p}) = c^{Z_m}(\mu, \nu)$$

and, under some conditions on either μ or ν, the maximizer is unique up to a uniform translation $\vec{p} \to \vec{p} + \beta(1, \ldots 1)$ on $\mathbb{R}^m$. Moreover, the maximizers of Ξ^ν_{μ,Z_m} yield the *unique* partitions $\{A_z, B_z;\ z \in Z_m\}$ of (8).

The accuracy of the approximation of $c(x, y)$ by $c^{Z_m}(x, y)$ depends, of course, on the choice of the set Z_m. In the special (but interesting) case $X = Y = Z = \mathbb{R}^d$ and $c(x, y) = |x - y|^\sigma$, $\sigma > 1$ it can be shown that $c^{Z_m}(x, y) - c(x, y) = O(m^{-2/d})$ for any x, y in a compact set, where Z_m are distributed on a regular grid containing this set.

From (7) and the above reasoning we obtain in particular

$$c^{Z_m}(\mu, \nu) - c(\mu, \nu) \geq 0 \tag{9}$$

for any pair of probability measures, and that, for a reasonable choice of Z_m, (9) is of order $m^{-2/d}$ if the supports of μ, ν are contained in a compact set.

For a given $m \in \mathbb{N}$ and pair of probability measures μ, ν and , the optimal choice of Z_m is the one which minimizes (9). Let

$$\phi^m(\mu, \nu) := \inf_{Z_m \subset Z} c^{Z_m}(\mu, \nu) - c(\mu, \nu) \geq 0 \tag{10}$$

where the infimum is over all sets of m points in Z. Note that the optimal choice now depends on the measures μ, ν themselves (and not only on their supports). A natural question is then to evaluate the assymptotic limits

$$\bar{\phi}(\mu, \nu) := \limsup_{m \to \infty} m^{2/d} \underline{c}^m(\mu, \nu) \quad ; \quad \underline{\phi}(\mu, \nu) := \liminf_{m \to \infty} m^{2/d} \underline{c}^m(\mu, \nu) \ .$$

Some preliminary results regarding these limits are discussed in this paper.

1.3 Numerical method

The numerical calculation of (3) we advertise in this paper apply the semi-discrete approximation c^{Z_m} of order m. It also involves discretization of μ, ν into atomic measures of finite support (n). The level of approximation is determined by the two parameters: The cardinality of the supports of the discretized measures, n, and the cardinality of the semi-finite approximation m of the cost. The idea of semi-discrete approximation is to choose n much larger than m. As we shall see, the evaluation of the approximate solution involves finding a maximizer to a concave function in m variables, where the complexity of calculating this function, and each of its partial derivatives, is of order n. A naive gradient descent method then result in $O(m)$ iterations to approximate this maximum, where each iteration is of order mn. This yields a complexity of order $O(m^2 n)$ to obtain a transport plan on the approximation level of $m^{-2/d}$. This should be compared to the n^3 complexity of the Hungarian algorithm [17]. We shall not, however, pursue a rigorous complexity estimate in this paper.

1.4 Structure of the paper

In section 2 we consider optimal partitions in the weak sense of probability measures, as Kantorovich relaxation of solutions of the optimal transport in semi-discrete setting. We formulate and prove a duality theorem (Theorem 2.1) which yields the relation between the minimizer of the OMT with semi-discrete cost to maximizing a dual function Ξ of m variables.

In section 3 we define strong partitions of the domains, and introduce conditions for the uniqueness of optimal solution and its representation as the analogue of optimal Monge mapping. The main results of this section is given in Theorem 3.1. In section 4 we introduce an interesting application of this concept to the theory of pricing of goods in Hedonic markets, and remark on possible generalization of optimal partitions to *optimal subpartition*. This model, related generalizations and further analysis will be pursued in a separate publication.

In section 5 we discuss optimal sampling of fixed number of centers (m). In particular we show a monotone sequence of improving semi-discrete approximation by floating the m centers into improved positions. In section 5.2 we provide some assymptotic properties of the error of the semi-discrete approximation as $m \to \infty$.

In section 6 we introduces a detailed description of the algorithm on the discrete level.

In section 7 we show some numerical experiments of calculating optimal partitions in the case of quadratic cost functions on a planar domain.

The numerical method we propose in this paper has some common features with the approach of Merigot [14], see also [4], as we recently discovered. We shall discuss this issues in section 8.

1.5 Notations and standing assumptions

1. X, Y are Polish (complete, separable) metric spaces.
2. $\mathcal{M}_+(X)$ is the cone of non-negative Borel measures on X (resp. for Y).
3. The $weak-*$ topology on $\mathcal{M}_+(X)$ is the dual of $C_b(X)$, the space of bounded continuous functions on X (resp. for Y).
4. $\mathcal{M}_1(X)$ is the cone of probability (normalized) non-negative Borel measures in $\mathcal{M}_+(X)$ (resp. for Y).
5. For $\mu \in \mathcal{M}_1(X)$, $\nu \in \mathcal{M}_1(Y)$,
   $\Pi_X^Y(\mu, \nu) := \{\pi \in \mathcal{M}_1(X \times Y) \ ; \mu \text{ is the } X \text{ marginal and } \nu \text{ is the } Y \text{ marginal of } \ \pi\}$
6. The $m-$simplex $\Sigma_m := \{\vec{s} := (s_1, \dots s_m), \ s_i \geq 0, \ \sum_{i=1}^m s_i = 1\} \subset \mathbb{R}^m$.

2 Optimal partitions

Definition 2.1.

i) A $m-$partition of a pair of a probability measure $\mu \in \mathcal{M}_1(X)$ subjected to $\vec{r} \in \Sigma_m$ is given by m nonnegative measures $\mu_z \in \mathcal{M}_+(X)$ on X such that $\sum_{z \in Z_m} \mu_z = \mu$ and $\int_X d\mu_z = r_z$. The set of all such partitions $\vec{\mu} := (\mu_1, \dots \mu_m)$ is denoted by $\mathcal{P}_X^{\vec{r}}(\mu)$.

ii) If, in addition, $\nu \in \mathcal{M}_1(Y)$ then $(\vec{\mu}, \vec{\nu}) \in \mathcal{P}_X^Y(\mu, \nu)$ iff $\vec{\mu} \in \mathcal{P}_X^{\vec{r}}(\mu)$ and $\vec{\nu} \in \mathcal{P}_Y^{\vec{r}}(\nu)$ for some *$\vec{r} \in \Sigma_m$.*

The following Lemma is a result of compactness of probability Borel measure on a compact space (see e.g. [5]).

Lemma 2.1. *For any $\vec{r} \in \Sigma_m$, the set of partitions $\mathcal{P}_X^{\vec{r}}$ is compact with respect to the $(C^*)^m(X)$ topology. In addition, $\mathcal{P}_X^Y(\mu, \nu)$ is compact with respect to $(C^*)^m(X) \times (C^*)^m(Y)$ topology.*

Lemma 2.2.

$$c^{Z_m}(\mu, \nu) = \min_{(\vec{\mu}, \vec{\nu}) \in \mathcal{P}_X^Y(\mu, \nu)} \sum_{z \in Z_m} \left[\int_X c^{(1)}(x, z) \mu_z(dx) + \int_Y c^{(2)}(z, y) \nu_z(dy) \right]$$

where $c^{Z_m}(\mu, \nu)$ as defined by (3, 7) and $(\vec{\mu}, \vec{\nu}) \in \mathcal{P}_X^Y(\mu, \nu)$.

Proof. First note that the existence of minimizer is obtained by Lemma 2.1.

Define, for $z \in Z_m$,

$$\Gamma_z := \{(x,y) \in X \times Y;\ c^{(1)}(x,z) + c^{(2)}(z,y) \le c^{Z_m}(x,y)\} \subset X \times Y$$

such that Γ_z is measurable in $X \times Y$, $\Gamma_z \cap \Gamma_{z'} = \emptyset$ if $z \neq z^{'}$ and $\sum_{z\in Z_m} \Gamma_z = X \times Y$. Note that, in general, the choice of $\{\Gamma_z\}$ is not unique.

Given $\pi \in \Pi_X^Y(\mu,\nu)$, let π_z be the restriction of π to Γ_z. In particular $\sum_{z\in Z_m} \pi_z = \pi$. Let μ_z be the X marginal of π_z and ν_z the y marginal of π_z. Then $(\vec{\mu},\vec{\nu})$ defined in this way is in $\mathcal{P}_X^Y(\mu,\nu)$. Since by definition $c^{Z_m}(x,y) = c^{(1)}(x,z) + c^{(2)}(z,y)$ a.s. π_z,

$$\begin{aligned}\int_X \int_Y c^{Z_m}(x,y)\pi(dxdy) &= \sum_{z\in Z_m} \int_X \int_Y c^{Z_m}(x,y)\pi_z(dxdy) \\ &= \sum_{z\in Z_m} \int_X \int_Y (c^{(1)}(x,z)\pi_z(dxdy) + \int_X (c^{(2)}(z,y)\pi_z(dxdy)\pi_z(dxdy) \\ &= \sum_{z\in Z_m} \left[\int_X c^{(1)}(x,z)\mu_z(dx) + \int_Y c^{(2)}(z,y)\nu_z(dy) \right] \quad (11)\end{aligned}$$

Choosing π above to be the optimal transport plan we get the inequality

$$c^{Z_m}(\mu,\nu) \ge \inf_{(\vec{\mu},\vec{\nu})\in\mathcal{P}_X^Y(\mu,\nu)} \sum_{z\in Z_m} \left[\int_X c^{(1)}(x,z)\mu_z(dx) + \int_Y c^{(2)}(z,y)\nu_z(dy) \right] .$$

To obtain the opposite inequality, let $(\vec{\mu},\vec{\nu}) \in \mathcal{P}_X^Y(\mu,\nu)$ and set $r_z := \int_X d\mu_z \equiv \int_Y d\nu_z$. Define $\pi(dxdy) = \sum_{z\in Z_m} r_z^{-1}\mu_z(dx)\nu_z(dy)$. Then $\pi \in \Pi_X^Y(\mu,\nu)$ and, from (7)

$$\begin{aligned}\int_X \int_Y c^{Z_m}(x,y)\pi(dxdy) &= \sum_{z\in Z_m} \int_X \int_Y c^{Z_m}(x,y)r_z^{-1}\mu_z(dx)\nu_z(dy) \\ &\le \sum_{z\in Z_m} \int_X (c^{(1)}(x,z) + c^{(2)}(z,y))r_z^{-1}\mu_z(dx)\nu_z(dy) \\ &= \sum_{z\in Z_m} \left[\int_X c^{(1)}(x,z)\mu_z(dx) + \int_Y c^{(2)}(z,y)\nu_z(dy) \right] \quad (12)\end{aligned}$$

and we get the second inequality. $\square$

Given $\vec{p} = (p_{z_1}, \dots p_{z_m}) \in \mathbb{R}^m$, let

$$\xi^{(1)}_{Z_m}(\vec{p},x) := \min_{z\in Z_m} c^{(1)}(x,z) + p_z \quad ; \quad \xi^{(2)}_{Z_m}(\vec{p},y) := \min_{z\in Z_m} c^{(2)}(z,y) + p_z \tag{13}$$

$$\Xi_\mu^{Z_m}(\vec{p}) := \int_X \xi^{(1)}_{Z_m}(\vec{p},x)\mu(dx) \quad ; \quad \Xi_\nu^{Z_m}(\vec{p}) := \int_Y \xi^{(2)}_{Z_m}(\vec{p},y)\nu(dy) \ . \tag{14}$$

$$\Xi^\nu_{\mu,Z_m}(\vec{p}) := \Xi_\mu^{Z_m}(\vec{p}) + \Xi_\nu^{Z_m}(-\vec{p}) \ . \tag{15}$$

Lemma 2.3. *If $\mu \in \mathcal{M}_1(X)$ then for any $\vec{r} \in \Sigma_m$,*

$$(-\Xi_\mu^{Z_m})^*(-\vec{r}) := \sup_{\vec{p}\in\mathbb{R}^m} \Xi_\mu^{Z_m}(\vec{p}) - \vec{p}\cdot\vec{r} = c^{(1)}\left(\mu, \sum_{z\in Z_m} r_z\delta_z\right) = \min_{\vec{\mu}\in\mathcal{P}_X^{\vec{r}}(\mu)} \sum_{z\in Z_m} \int_X c^{(1)}(x,z)\mu_z(dx) \ . \tag{16}$$

Analogously, for $\nu \in \mathcal{M}_1(Y)$

$$(-\Xi_\nu^{Z_m})^*(-\vec{r}) := \sup_{\vec{p}\in\mathbb{R}^m} \Xi_\nu^{Z_m}(\vec{p}) - \vec{p}\cdot\vec{r} = c^{(2)}\left(\nu, \sum_{z\in Z_m}\delta_z\right) = \min_{\vec{\nu}\in\mathcal{P}_Y^{\vec{r}}(\nu)} \sum_{z\in Z_m}\int_Y c^{(2)}(z,y)\nu_z(dy) \ . \quad (17)$$

Here $\vec{p}\cdot\vec{r} := \sum_{z\in Z_m} r_z p_z$.

Proof. This is a special case of the general duality theorem of Monge-Kantorovich. See, for example [22]. It is also a special case of generalized partitions, see Theorem 3.1 and its proof in [24].

□

Theorem 2.1.

$$\sup_{\vec{p}\in\mathbb{R}^m} \Xi_{\mu,Z_m}^{\nu}(\vec{p}) = c^{Z_m}(\mu,\nu) \ . \quad (18)$$

Proof. From Lemma 2.2, Lemma 2.3 and Definition 2.1 we obtain

$$c^{Z_m}(\mu,\nu) = \inf_{\vec{r}\in\Sigma_m}\left[(-\Xi_\mu^{Z_m})^*(-\vec{r}) + (-\Xi_\nu^{Z_m})^*(-\vec{r})\right] \ . \quad (19)$$

Note that $(-\Xi_\mu^{Z_m})^*$, $(-\Xi_\nu^{Z_m})^*$ as defined in (16, 17), are, in fact, the Legendre transforms of $-\Xi_\mu^{Z_m}$, $-\Xi_\nu^{Z_m}$, respectively. As such, they are defined formally on the whole domain $\mathbb{R}^m$ (considered as the dual of itself under the canonical inner product). It follows that $(-\Xi_\mu^{Z_m})^*(\vec{r}) = (-\Xi_\nu^{Z_m})^*(\vec{r}) = \infty$ for $\vec{r}\in\mathbb{R}^m - \Sigma_m$. Note that this definition is consistent with the right hand side of (16, 17), since $\mathcal{P}_X^{\vec{r}}(\mu) = \mathcal{P}_Y^{\vec{r}}(\nu) = \emptyset$ for $\vec{r}\notin\Sigma_m$.

On the other hand, $\Xi_\mu^{Z_m}$ and $\Xi_\nu^{Z_m}$ are both finite and continuous on the whole of $\mathbb{R}^m$. The Fenchel-Rockafellar duality theorem (see [22]- Thm 1.9) then implies

$$\sup_{\vec{p}\in\mathbb{R}^m} \Xi_\mu^{Z_m}(\vec{p}) + \Xi_\nu^{Z_m}(-\vec{p}) = \inf_{\vec{r}\in\mathbb{R}^m}(-\Xi_\mu^{Z_m})^*(\vec{r}) + (-\Xi_\nu^{Z_m})^*(\vec{r}) \ . \quad (20)$$

The proof follows from (15, 19).

An alternative proof:
We can prove (18) directly by constrained minimization, as follows: $(\vec{\mu},\vec{\nu})\in\mathcal{P}_X^Y(\mu,\nu)$ iff $F(\vec{p},\phi,\psi) :=$

$$\sum_{z\in Z_m} p_i\left(\int_X d\mu_i - \int_Y d\nu_i\right) + \int_X \phi(x)\left(\mu(dx) - \sum_{z\in Z_m}\mu_z(dx)\right) + \int_Y \psi(y)\left(\nu(dy) - \sum_{z\in Z_m}\nu_z(dy)\right) \le 0$$

for any choice of $\vec{p}\in\mathbb{R}^m$, $\phi\in C(X)$, $\psi\in C(Y)$. Moreover, $\sup_{\vec{p},\phi,\psi} F = \infty$ unless $(\vec{\mu},\vec{\nu})\in\mathcal{P}_X^Y(\mu,\nu)$. We can then obtain from Lemma 2.2: $c^{Z_m}(\mu,\nu) =$

$$\inf_{\{\mu_z\in\mathcal{M}_+(X),\nu_z\in\mathcal{M}_+(Y)\}} \sup_{\vec{p}\in\mathbb{R}^m,\phi\in C(X),\psi\in C(Y)} \sum_{z\in Z_m}\left[\int_X c^{(1)}(x,z)\mu_z(dx) + \int_Y c^{(2)}(z,y)\nu_z(dy)\right] + F(\vec{p},\phi,\psi)$$

$$= \sup_{\vec{p}\in\mathbb{R}^m,\phi\in C(X),\psi\in C(Y)} \inf_{\{\mu_z\in\mathcal{M}_+(X),\nu_z\in\mathcal{M}_+(Y)\}} \sum_{z\in Z_m}\int_X\left(c^{(1)}(x,z) + p_z - \phi(x)\right)\mu_z(dx)$$

$$+ \sum_{z\in Z_m}\int_Y\left(c^{(2)}(z,y) - p_z - \psi(y)\right)\nu_z(dy) + \int_X \phi\mu(dx) + \int_Y \psi\nu(dy) \ . \quad (21)$$

We now observe that the infimum on $\{\mu_z, \nu_z\}$ above is $-\infty$ unless $c^{(1)}(x,z) + p_z - \phi(x) \geq 0$ and $c^{(2)}(z,y) + p_z - \psi(y) \geq 0$ for any $z \in Z_m$. Hence, the two sums on the right of (21) are non-negative, so the infimum with respect to $\{\mu_z, \nu_z\}$ is zero. To obtain the supremum on the last two integrals on the right of (21) we choose ϕ, ψ as large as possible under this constraint, namely

$$\phi(x) = \min_{z \in Z_m} c^{(1)}(x,z) + p_z \quad , \quad \psi(y) = \min_{z \in Z_m} c^{(2)}(z,y) - p_z$$

so $\phi(x) \equiv \xi^{(1)}_{Z_m}(\vec{p}, x)$, $\psi(y) \equiv \xi^{(2)}_{Z_m}(-\vec{p}, y)$ by definition via (13). $\square$

3 Strong partitions

We now define strong partitions as a special case of partitions (Definition 2.1).

Definition 3.1.

i) A partition $\vec{\mu} \in \mathcal{P}^{\vec{r}}_X(\mu)$ is called a strong $m-$partition *if there exists m measurable sets $A_z \subset X$, $z \in Z_m$ which are essentially disjoint, namely $\mu(A_z \cap A_{z'}) = \emptyset$ for $z \neq z'$ and $\mu(\cup_{z \in Z_m} A_z) = X$, such that μ_z is the restriction of μ to A_z. The set of strong $m-$partition corresponding to $\vec{r} \in \Sigma_m$ is denoted by $\widehat{\mathcal{P}}^{\vec{r}}_X(\mu)$.*

ii) In addition, for $\nu \in \mathcal{M}_1(Y)$ then $(\vec{\mu}, \vec{\nu}) \in \widehat{\mathcal{P}}^{Y}_X(\mu, \nu)$ iff $\vec{\mu} \in \widehat{\mathcal{P}}^{\vec{r}}_X(\mu)$ and $\vec{\nu} \in \widehat{\mathcal{P}}^{\vec{r}}_Y(\nu)$ for some *$\vec{r} \in \Sigma_m$. In particular, a strong $m-$partition is composed of m μ measurable sets $A_z \subset X$ and m ν measurable sets $B_z \subset Y$ such that $\int_{A_z} d\mu = \int_{B_z} d\nu$ for $z \in Z_m$.*

Assumption 3.1. .

a) $\mu \in \mathcal{M}_1(X)$ is atomless and $\mu(x; c^{(1)}(x,z) - c^{(1)}(x,z') = p) = 0$ for any $p \in \mathbb{R}$ and any $z, z' \in Z_m$.

b) $\nu \in \mathcal{M}_1(Y)$ is atomless and $\nu(y; c^{(2)}(z,y) - c^{(2)}(z',y) = p) = 0$ for any $p \in \mathbb{R}$ and any $z, z' \in Z_m$.

Let us also define, for $\vec{p} \in \mathbb{R}^m$

$$A_z(\vec{p}) := \{x \in X;\ c^{(1)}(x,z) + p_z = \xi^{(1)}_{Z_m}(\vec{p}, x)\} \quad ; \quad B_z(\vec{p}) := \{y \in Y;\ c^{(2)}(z,y) + p_z = \xi^{(2)}_{Z_m}(\vec{p}, y)\} \ . \tag{22}$$

Note that, by (13, 14)

$$\Xi^{Z_m}_{\mu}(\vec{p}) = \sum_{z \in Z_m} \int_{A_z(\vec{p})} (c^{(1)}(x,z) + p_z)\mu(dx) \tag{23}$$

likewise

$$\Xi^{Z_m}_{\nu}(\vec{p}) = \sum_{z \in Z_m} \int_{B_z(\vec{p})} (c^{(2)}(z,y) + p_z)\nu(dy) \ . \tag{24}$$

Lemma 3.1. *Under assumption 3.1 (a) (resp. (b))*

i) For any $\vec{p} \in \mathbb{R}^m$, $\{A_z(\vec{p})\}$ (resp. $\{B_z(\vec{p})\}$) induces essentially disjoint partitions of X (resp. Y).

ii) $\Xi_\mu^{Z_m}$ *(resp.* $\Xi_\nu^{Z_m}$*) is continually differentiable functions on* $\mathbb{R}^m$*,*

$$\frac{\partial \Xi_\mu^{Z_m}}{\partial p_z} = \mu(A_z(\vec{p})) \quad \textit{resp.} \quad \frac{\partial \Xi_\nu^{Z_m}}{\partial p_z} = \nu(B_z(\vec{p})) \ .$$

This Lemma is a special case of Lemma 4.3 in [W].

Theorem 3.1. *Under either assumption 3.1-(a) or (b) there exists a unique minimizer* $\vec{r}_0$ *of (19). In addition, there exists a maximizer* $\vec{p}_0 \in \mathbb{R}^m$ *of* Ξ^ν_{μ,Z_m}*, and either (in case (a))* $\{A_z(\vec{p}_0)\}$ *or (in case (b))* $\{B_z(-\vec{p}_0)\}$ *induces a corresponding strong* $m-$*partition in (a)* $\widehat{\mathcal{P}}_X^{\vec{r}_0}(\mu)$ *or (b)* $\widehat{\mathcal{P}}_Y^{\vec{r}_0}(\nu)$*. In particular, if* both *(a+b) holds then* $\{A_z(\vec{p}_0), \{B_z(-\vec{p}_0)\}$ *induces a strong* $m-$*partition in* $\widehat{\mathcal{P}}_X^Y(\mu,\nu)$*, and*

$$\pi_0(dxdy) := \sum_{z\in Z_m; r_{0,z}=\mu(A_z(\vec{p}_0))} (r_{0,z})^{-1}\mathbf{1}_{A_z(\vec{p}_0)}(x)\mathbf{1}_{B_z(-\vec{p}_0)}(y)\mu(dx)\nu(dy) \tag{25}$$

is the unique optimal transport plan for $c^{Z_m}(\mu,\nu)$*.*

Proof. Note that $\Xi(\vec{p})-\vec{r}\cdot\vec{p}$ is invariant under additive shift for $\Xi = \Xi_\mu^{Z_m}, \Xi_\nu^{Z_m}$ and $\vec{r}\in\Sigma_m$. Indeed, $\Xi(\vec{p}+\alpha\vec{1}) = \Xi(\vec{p})+\alpha$ for any $\alpha\in\mathbb{R}$ where $\vec{1} := (1,\ldots 1)$. So, we restrict the domain of Ξ to

$$\vec{p}\in R^m \quad , \quad \vec{p}\cdot\vec{1} = 0 \ . \tag{26}$$

Assume (a). Given $\vec{r}\in\Sigma_m$. Assume first

$$r_z \in (0,1) \quad \text{for any } z\in Z_m \ . \tag{27}$$

We prove the existence of a maximizer $\vec{p}_0$,

$$(-\Xi_\mu^{Z_m})^*(-\vec{r}) = \Xi_\mu^{Z_m}(\vec{p}_0) - \vec{p}_0\cdot\vec{r} \geq \Xi_\mu^{Z_m}(\vec{p}) - \vec{p}\cdot\vec{r}$$

for any $\vec{p}\in\mathbb{R}^m$. Let $\vec{p}_n$ be a maximizing sequence, that is

$$\lim_{n\to\infty} \Xi_\mu^{Z_m}(\vec{p}_n) - \vec{p}_n\cdot\vec{r} = (-\Xi_\mu^{Z_m})^*(-\vec{r})$$

(c.f. (17)).

Let $\|\vec{p}\|_2 := (\sum_{z\in Z_m} p_z^2)^{1/2}$ be the Euclidian norm of $\vec{p} = (p_{z_1},\ldots p_{z_m}) \in \mathbb{R}^m$. If we prove that for any maximizing sequence $\vec{p}_n$ the norms $\|\vec{p}_n\|_2$ are uniformly bounded, then there exists a converging subsequence whose limit is the maximizer $\vec{p}_0$. This follows, in particular, since $\Xi_\mu^{Z_m}$ is a closed (upper-semi-continuous) function.

Assume there exists a subsequence along which $\|\vec{p}_n\|_2\to\infty$. Let $\hat{\vec{p}}_n := \vec{p}_n/\|\vec{p}_n\|_2$. Let

$$\begin{aligned}\Xi_\mu^{Z_m}(\vec{p}_n) - \vec{p}_n\cdot\vec{r} &:= \left[\Xi_\mu^{Z_m}(\vec{p}_n) - \vec{p}_n\cdot\nabla_{\vec{p}}\Xi_\mu^{Z_m}(\vec{p}_n)\right] + \vec{p}_n\cdot\left(\nabla_{\vec{p}}\Xi_\mu^{Z_m}(\vec{p}_n)-\vec{r}\right)\\ &= \left[\Xi_\mu^{Z_m}(\vec{p}_n) - \vec{p}_n\cdot\nabla_{\vec{p}}\Xi_\mu^{Z_m}(\vec{p}_n)\right] + \|\vec{p}_n\|_2\hat{\vec{p}}_n\cdot\left(\nabla_{\vec{p}}\Xi_\mu^{Z_m}(\vec{p}_n)-\vec{r}\right) \ .\end{aligned} \tag{28}$$

In addition, by (23) and Lemma3.1-(ii)

$$\begin{aligned}-\infty < \int_X \min_{z\in Z_m} c^{(1)}(x,z)\mu(dx) \leq \left[\Xi_\mu^{Z_m}(\vec{p}) - \vec{p}\cdot\nabla_{\vec{p}}\Xi_\mu^{Z_m}(\vec{p})\right] &= \sum_{z\in Z_m}\int_{A_z(\vec{p})} c^{(1)}(x,z)\mu(dx)\\ &\leq \int_X \max_{z\in Z_m} c^{(1)}(x,z)\mu(dx) < \infty \ .\end{aligned} \tag{29}$$

By (28- 29) we obtain, for $\|\vec{p}_n\|_2 \to \infty$,

$$\lim_{n\to\infty} \hat{\vec{p}}_n \cdot \left(\nabla_{\vec{p}}\Xi_\mu^{Z_m}(\vec{p}_n) - \vec{r}\right) = 0 \ . \tag{30}$$

Since $\hat{\vec{p}}_n$ lives in the unit sphere S^{m-1} in $\mathbb{R}^m$ (which is a compact set), there exists a subsequence for which $\hat{\vec{p}}_n \to \hat{\vec{p}}_0 := (\hat{p}_{0,z_1}, \ldots \hat{p}_{0,z_m}) \in S^{m-1}$. Let $P_- := \min_{z\in Z_m} \hat{p}_{z,0}$ and $J_- := \{z \in Z_m \ ; \hat{p}_{0,z} = P_-\}$.

Note that for $n \to \infty$ along such a subsequence, $p_{n,z} - p_{n,z'} \to -\infty$ for $z \in J_-, z' \notin J_-$. It follows that $A_{z'}(\vec{p}_n) = \emptyset$ if $z' \notin J_-$ for n large enough, hence $\cup_{z\in J_-} A_z(\vec{p}_n) = X$ for n large enough. Let μ_z^n be the restriction of μ to $A_z(\vec{p}_n)$. Then the limit $\mu_z^n \rightharpoonup \mu_z$ exists (along a subsequence) where $n \to \infty$. In particular, by Lemma 3.1

$$\lim_{n\to\infty} \frac{\partial \Xi_\mu^{Z_m}}{\partial p_{n,z}}(\vec{p}_n) = \int_X d\mu_z$$

while $\mu_z \neq 0$ if only if $z \in J_-$, and $\sum_{z\in J_-} \mu_z = \mu$. Since $\hat{\vec{p}}_{0,z} = P_-$ for $z \in J_-$ is the minimal value of the coordinates of $\hat{\vec{p}}_0$, it follows that

$$\lim_{n\to\infty} \hat{\vec{p}}_n \cdot \left(\nabla_{\vec{p}}\Xi_\mu^{Z_m}(\vec{p}_n) - \vec{r}\right) = -\vec{r}\cdot\hat{\vec{p}}_0 + P_- \sum_{z\in J_-} \int_X d\mu_z = -\vec{r}\cdot\hat{\vec{p}}_0 + P_- \ .$$

Now, by (27), $\vec{r}\cdot\vec{p}_0 > P_-$ unless $J_- = Z_m$. In the last case we obtain a contradiction of (26) since it implies $\hat{\vec{p}}_0 = 0$ which contradicts $\hat{\vec{p}}_0 \in S^{m-1}$. If J_- is a proper subset of Z_m we obtain a contradiction to (30).

If (27) is violated we may restrict to domain of $\Xi_\mu^{Z_m}$ to a subspace by eliminating all coordinates $z \in Z_m$ for which $r_z = 0$. On the restricted subspace we have a minimizer $\vec{p}_0$ by the above proof. Then we may extend $\vec{p}_0$ by assigning p_z sufficiently small if $r_z = 0$. This guarantees $A_z(\vec{p}_0) = \emptyset$, hence (Lemma 3.1) $\partial\Xi_\mu^{Z_m}/\partial p_z = 0$ for any such z. Hence the extended $\vec{p}_0$ is still a critical point of $\Xi_\mu^{Z_m}(\vec{p}) - \vec{r}\cdot\vec{p}$, and is a maximizer by concavity of $\Xi_\mu^{Z_m}$.

Next, we prove that $A_z(\vec{p}_0)$ is a unique optimal partition of X. Let $\vec{\mu} \in \mathcal{P}_X^r$ be a minimizer of (16). Since $\int_X d\mu_z = r_z$, $\sum_{z\in Z_m} \mu_z = \mu$, (16) implies

$$\sum_{z\in Z_m} \int_X c^{(1)}(x,z)\mu_z(dx) = (-\Xi_\mu^{Z_m})^*(-\vec{r})$$

and

$$(-\Xi_\mu^{Z_m})^*(-\vec{r}) = \Xi_\mu^{Z_m}(\vec{p}_0) - \vec{r}\cdot\vec{p}_0 = \int_X \xi_{Z_m}^{(1)}(\vec{p}_0, x)d\mu - \vec{p}_0\cdot\vec{r} = \sum_{z\in Z_m} \int_X \left(\xi_{Z_m}^{(1)}(\vec{p}_0, x) - p_{0,z}\right) d\mu_z \ ,$$

so

$$\sum_{z\in Z_m} \int_X \left(\xi_{Z_m}^{(1)}(\vec{p}_0, x) - p_{0,z} - c^{(1)}(x,z)\right) \mu_z(dx) = 0 \ .$$

On the other hand, $\xi_{Z_m}^{(1)}(\vec{p}_0, x) - p_{0,z} - c^{(1)}(x,z) \leq 0$ for any $x \in X$ by definition (13), so we must have the equality

$$\xi_{Z_m}^{(1)}(\vec{p}_0, x) = p_{0,z} + c^{(1)}(x,z)$$

a.e. on $\mathrm{supp}(\mu_z)$. Hence $\mathrm{supp}(\mu_z) \subset A_z(\vec{p}_0)$. Since $A_z(\vec{p}_0)$ are mutually disjoint and $\sum_{z\in Z_m} \mu_z = \mu$, then μ_z is necessarily the restriction of μ to $A_z(\vec{p}_0)$. On the other hand, for any $\vec{p} \neq \vec{p}_0 \mod \mathbb{R}\vec{1}$ there exists $z \in Z_m$ for which $\mu\left(A_z(\vec{p}_0)\Delta A_z(\tilde{\vec{p}})\right) \neq 0$. This implies that the strong partition $\vec{A}(\vec{p}_0)$ is the unique one.

The same result is applied to $\Xi_\nu^{Z_m}(\vec{p}) - \vec{p}_0 \cdot \vec{r}$. If we show that the minimizer $\vec{r}_0$ of the right side of (20) is unique, then it follows that the maximizer $\vec{p}_0$ of the left side of (20) is unique as well (up to $\vec{1}\mathbb{R}$), and, in particular, the optimal partition is unique. Hence, we only have to show the uniqueness of the minimizer of the right side of (20). This, in turn, follows if either $(-\Xi_\mu^{Z_m})^*$ or $(-\Xi_\nu^{Z_m})^*$ is *strictly convex*.

To prove this we recall some basic elements form convexity theory (see, e.g. [BC]):

i) If F is a convex function on $\mathbb{R}^m$ (say), then the sub gradient ∂F at point $p \in \mathbb{R}^m$ is defined as follows: $\vec{q} \in \partial F(\vec{p})$ if and only if

$$F(\vec{p'}) - F(\vec{p}) \geq \vec{q} \cdot (\vec{p'} - \vec{p}) \quad \forall \vec{p'} \in \mathbb{R}^m \ .$$

ii) The Legendre transform of F:

$$F^*(\vec{q}) := \sup_{\vec{p}\in\mathbb{R}^m} \vec{p} \cdot \vec{q} - F(\vec{p}) \ ,$$

and $Dom(F^*) \subset \mathbb{R}^m$ is the set on which $F^* < \infty$.

iii) The function F^* is convex (and closed), but $Dom(F^*)$ can be a proper subset of $\mathbb{R}^m$ (or even an empty set).

iv) The subgradient of a convex function is non-empty (and convex) at any point in the proper domain of this function (i.e. at any point in which the function takes a value in $\mathbb{R}$).

v) Young's inequality

$$F(\vec{p}) + F^*(\vec{q}) \geq \vec{p} \cdot \vec{q}$$

holds for any pair of points $(\vec{p}, \vec{q}) \in \mathbb{R}^m \times \mathbb{R}^m$. The equality holds iff $\vec{q} \in \partial F(\vec{p})$, iff $\vec{p} \in \partial F^*(\vec{q})$.

vi) The Legendre transform is involuting, i.e $F^{**} = F$ if F is convex and closed.

vii) A convex function is continuously differentiable in the interior of its proper domain iff its subgradient at any point in the interior of its domain is a singleton.

Returning to our case, let $F := -\Xi_\mu^{Z_m}$. It is a closed, convex, proper and continuously differentiable function defined everywhere on $\mathbb{R}^m$. Assume $(-\Xi_\mu^{Z_m})^*$ is not strictly convex. It means there exists $\vec{r}_1 \neq \vec{r}_2 \in Dom(-\Xi_\mu^{Z_m})^*$ for which

$$(-\Xi_\mu^{Z_m})^*(\frac{\vec{r}_1 + \vec{r}_2}{2}) = \frac{(-\Xi_\mu^{Z_m})^*(\vec{r}_1) + (-\Xi_\mu^{Z_m})^*(\vec{r}_2)}{2} \ . \tag{31}$$

Let $\vec{r} := \vec{r}_1/2 + \vec{r}_2/2$, and $\vec{p} \in \partial(-\Xi_\mu^{Z_m})^*(\vec{r})$. Then, by (iv, v)

$$0 = (-\Xi_\mu^{Z_m})^*(\vec{r}) + (-\Xi_\mu^{Z_m})^{**}(\vec{p}) - \vec{p} \cdot \vec{r} = (-\Xi_\mu^{Z_m})^*(\vec{r}) - \Xi_\mu^{Z_m}(\vec{p}) - \vec{p} \cdot \vec{r} \ . \tag{32}$$

By (31, 32):

$$\frac{1}{2}\left((-\Xi_\mu^{Z_m})^*(\vec{r}_1) - \Xi_\mu^{Z_m}(\vec{p}) - \vec{p}\cdot\vec{r}_1\right) + \frac{1}{2}\left((-\Xi_\mu^{Z_m})^*(\vec{r}_2) - \Xi_\mu^{Z_m}(\vec{p}) - \vec{p}\cdot\vec{r}_2\right) = 0$$

while (v) also guarantees

$$(-\Xi_\mu^{Z_m})^*(\vec{r}_i) - \Xi_\mu^{Z_m}(\vec{p}) - \vec{p}\cdot\vec{r}_i \geq 0 \quad , \quad i = 1, 2 \ .$$

It follows

$$(-\Xi_\mu^{Z_m})^*(\vec{r}_i) - \Xi_\mu^{Z_m}(\vec{p}) - \vec{p}\cdot\vec{r}_i = 0 \quad , \quad i = 1, 2 \ ,$$

so, by (v) again, $\{\vec{r}_1, \vec{r}_2\} \in \partial\Xi_\mu^{Z_m}(\vec{p})$. This is a contradiction of (vii) since $\Xi_\mu^{Z_m}$ is continuously differentiable everywhere on $\mathbb{R}^m$ by Lemma 3.1.

Finally, we prove that π_0 given by (25) is an optimal plan. First observe that $\pi_0 \in \Pi(\mu, \nu)$, hence

$$c^{Z_m}(\mu, \nu) \leq \int_X \int_Y c^{Z_m}(x, y)\pi_0(dxdy) \ .$$

Then we get, from (7)

$$c^{Z_m}(\mu, \nu) \leq \int_X \int_Y c^{Z_m}(x, y)\pi_0(dxdy) \leq \sum_{z\in Z_m} \int_{A_z(\vec{p}_0)\times B_z(-\vec{p}_0)} (c^{(1)}(x, z)\mu(dx) + c_{(2)}(y, z)\nu(dy))$$

$$= \sum_{z\in Z_m} \left(\int_{A_z(\vec{p}_0)} c^{(1)}(x, z)\mu(dx) + \int_{B_z(-\vec{p}_0)} c^{(2)}(z, y)\nu(dy) \right) = \Xi(\vec{p}_0) \leq c^{Z_m}(\mu, \nu)$$

where the last equality from Theorem 2.1. In particular, the first inequality is an equality so π_0 is an optimal plan indeed. $\square$

4 Pricing in hedonic market

In adaptation to the model of Hedonic market [7] there are 3 components: The space of consumers (say, X), space of producers (say Y) and space of commodities, which we take here to be a finite set $Z_m := \{z_1, \ldots z_m\}$. The function $c^{(1)} := c^{(1)}(x, z)$ is the *negative* of the utility of commodity $z \in Z_m$ to consumer x, while $c^{(2)} := c^{(2)}(z, y)$ is the cost of producing commodity $z \in Z_m$ by the producer y.

Let μ be a probability measure on X representing the distribution of consumers, and ν a probability measure on Y representing the distribution of the producers. Following [7] we add the "null commodity" z_0 and assign the zero utility and cost $c^{(1)}(x, z_0) = c^{(2)}(z_0, y) \equiv 0$ on X (resp. Y). We understand the meaning that a consumer (producer) chooses the null commodity is that he/she avoids consuming (producing) any item from Z_m.

The object of pricing in Hedonic market is to find equilibrium prices for the commodities which will balance supply and demand: Given a price p_z for z, the consumer at x will buy the commodity z which minimize its loss $c^{(1)}(x, z) + p_z$, or will buy nothing (i.e. "buy" the null commodity z_0) if $\min_{z\in Z_m} c^{(1)}(x, z) + p_z > 0$), while producer at y will prefer to produce commodity z which maximize its profit $-c^{(2)}(z, y) + p_z$, or will produce nothing if $\max_{z\in Z_m} -c^{(2)}(z, y) + p_z < 0$. Using notation (13-15) we define

$$\xi_X^0(\vec{p}, x) := \min\{\xi_{Z_m}^{(1)}(\vec{p}, x), 0\} \quad ; \quad \xi_Y^0(\vec{p}, y) := \min\{\xi_{Z_m}^{(2)}(\vec{p}, y), 0\} \tag{33}$$

$$\Xi^0_\mu(\vec{p}) := \int_X \xi^0_X(\vec{p},x)\mu(dx) \quad ; \quad \Xi^0_\nu(\vec{p}) := \int_Y \xi^0_Y(\vec{p},y)\nu(dy) \ . \tag{34}$$

$$\Xi^{0,\nu}_\mu(\vec{p}) := \Xi^0_\mu(\vec{p}) + \Xi^0_\nu(-\vec{p}) \ . \tag{35}$$

Thus, $\Xi^{0,\nu}_\mu(\vec{p})$ is the difference between the total loss of all consumers and the total profit of all producers, given the prices vector $\vec{p}$. It follows that an equilibrium price vector balancing supply and demand is the one which (somewhat counter-intuitively) *maximizes* this difference. The corresponding optimal strong $m-$partition represent the matching between producers of $(B_z \subset Y)$ to consumers $(A_z \subset X)$ of $z \in Z$. The introduction of null commodity allows the possibility that only part of the consumer (producers) communities actually consume (produce), that is $\cup_{z\in Z_m} A_z \subset X$ and $\cup_{z\in Z_m} B_z \subset Y$, with $A_0 = X - \cup_{z\in Z_m} A_z$ $(B_0 = Y - \cup_{z\in Z_m} B_z)$ being the set of non-buyers (non-producers).

From the dual point of view, an adaptation $c_0^{Z_m}(x,y) := \min\{c^{Z_m}(x,y), 0\}$ of (7) (in the presence of null commodity) is the *cost of direct matching* between producer y and consumer x. The *optimal matching* (A_z, B_z) is the one which *minimizes* the total cost $c_0^{Z_m}(\mu,\nu)$ over all *sub-$m-$partitions* $\widehat{\mathcal{P}}^Y_X(\mu,\nu)$ as defined in Definition 3.1-(ii) with the possible inequality $\mu(\cup A_z) = \nu(\cup B_z) \leq 1$.

5 Dependence on the sampling set

So far we took the smapling set $Z_m \subset Z$ to be fixed. Here we consider the effect of optimizing Z_m within the sets of cardinality m in Z.

As we already know from (5, 7), $c^{Z_m}(x,y) \geq c(x,y)$ on $X\times Y$ for any $(x,y) \in X\times Y$ and $Z_m \subset Z$. Hence also $c^{Z_m}(\mu,\nu) \geq c(\mu,\nu)$ for any $\mu,\nu \in \mathcal{M}_1$ and any $Z_m \subset Z$ as well. An *improvement* of Z_m is a new choice $Z_m^{new} \subset Z$ of the *same* cardinality m such that $c^{Z_m^{new}}(\mu,\nu) < c^{Z_m}(\mu,\nu)$.

In section 5.1 we propose a way to improve a given $Z_m \subset Z$, once the optimal partition is calculated. Of course, the improvement depends on the measure μ,ν.

In section 5.2 we discuss the limit $m \to \infty$ and prove some assymptotic estimates.

5.1 Monotone improvement

Proposition 5.1. *Define Ξ^ν_{μ,Z_m} on $\mathbb{R}^m$ with respect to $Z_m := \{z_1,\dots z_m\} \in Z$ as in (15). Let $(\vec{\mu},\vec{\nu}) \in \mathcal{P}^Y_X(\mu,\nu)$ be the optimal partition corresponding to $c^{Z_m}(\mu,\nu)$. Let $\zeta(i) \in Z$ be a minimizer of*

$$Z \ni \zeta \mapsto \int_X c^{(1)}(x,\zeta)\mu_{z_i}(dx) + \int_Y c^{(2)}(\zeta,y)\nu_{z_i}(dy) \ . \tag{36}$$

Let $Z_m^{new} := \{\zeta(1),\dots\zeta(m)\}$. Then $c^{Z_m^{new}}(\mu,\nu) \leq c^{Z_m}(\mu,\nu)$.

Corollary 5.1. *Let Assumption 3.1 (a+b), and $\vec{p}_0$ be the minimizer of Ξ^{ν,Z_m}_μ in $\mathbb{R}^m$. Let $\{A_z(\vec{p}_0), B_z(-\vec{p}_0)\}$ be the strong partition corresponding to Z_m as in (22). Then the components of Z_m^{new} are obtained as the minimizers of*

$$Z \ni \zeta \mapsto \int_{A_z(\vec{p}_0)} c^{(1)}(x,\zeta)\mu(dx) + \int_{B_z(-\vec{p}_0)} c^{(2)}(\zeta,y)\nu(dy) \ .$$

Proof. (of Proposition 5.1): Let $\Xi_\mu^{\nu,new}$ be defined with respect to Z_m^{new}. By Lemma 2.2 and Theorem 2.1 $\Xi_\mu^{\nu,new}(\vec{p}) \leq \Xi_\mu^{\nu}(\vec{p}^*) := \max_{\mathbb{R}^m} \Xi_\mu^{\nu,Z_m}$ for any $\vec{p} \in \mathbb{R}^m$,
so $\max_{\mathbb{R}^m} \Xi_\mu^{\nu,new}(\vec{p}) \equiv c^{Z_m^{new}}(\mu,\nu) \leq \max_{\mathbb{R}^m} \Xi_\mu^{\nu,Z_m}(\vec{p}) \equiv c^{Z_m}(\mu,\nu)$. □

Remark 5.1. *If c is a quadratic cost then z^{new} is the center of mass of $A_z(\vec{p}_0)$ and $B_z(-\vec{p}_0)$:*

$$z^{new} := \frac{\int_{A_z(\vec{p}_0)} x\mu(dx) + \int_{B_z(-\vec{p}_0)} y\nu(dy)}{\mu(A_z(\vec{p}_0)) + \nu(B_z(-\vec{p}_0))} \ .$$

We shall take advantage of this in section 6.1.

Let

$$\underline{c}^m(\mu,\nu) := \inf_{Z_m \subset Z \ ; \ \#(Z_m)=m} c^{Z_m}(\mu,\nu) \ .$$

Let $Z_m^k := \{z_1^k, \ldots z_m^k\} \subset Z$ be a sequence of sets such that z_z^{k+1} is obtained from Z_m^k via (36). Then by Proposition 5.1

$$c(\mu,\nu) \leq \underline{c}^m(\mu,\nu) \leq \ldots c^{Z_m^{k+1}}(\mu,\nu) \leq c^{Z_m^k}(\mu,\nu) \leq \ldots c^{Z_m^0}(\mu,\nu) \ .$$

Open problem: Under which additional conditions one may gurantee

$$\lim_{k\to\infty} c^{Z_m^k}(\mu,\nu) = \underline{c}^m(\mu,\nu) \ \ ?$$

5.2 Assymptotic estimates

Recall the definition (10)

$$\phi^m(\mu,\nu) := \inf_{Z_m \subset Z} c^{Z_m}(\mu,\nu) - c(\mu,\nu) \geq 0 \ .$$

Consider the case $X = Y = Z = \mathbb{R}^d$ and

$$c(x,y) = \min_{z\in\mathbb{R}^d} h(|x-z|) + h(|y-z|)$$

where $h : \mathbb{R}^+ \to \mathbb{R}^+$ is convex, monotone increasing, twice continuous differentiable. Note that $c(x,y) = 2h(|x-y|/2)$.

Lemma 5.1. *Suppose both μ and ν are supported on in a compact set in $\mathbb{R}^d$. Then there exists $C = C(\mu,\nu) < \infty$ such that*

$$\limsup_{m\to\infty} m^{2/d}\phi^m(\mu,\nu) \leq C(\mu,\nu) \ . \tag{37}$$

Proof. By Taylor expansion of $z \to h(|x-z|) + h(|y-z|)$ at $z_0 = (x+y)/2$ we get

$$h(|x-z|) + h(|y-z|) = 2h(|x-y|/2) + \frac{1}{2|x-y|^2} h''\left(\frac{|x-y|}{2}\right)\left[(x-y)\cdot(z-z_0)\right]^2 + o^2(z-z_0) \ .$$

Let now Z_m be a regular grid of m points which contains the support K. The distance between any $z \in K$ to the nearest point in the grid does not exceed $C(K)m^{-1/d}$, for some constant $C(K)$.

Hence $c_m(x,y) - c(x,y) \le \sup|h^{''}|C(K)^2 m^{-2/d}$ if $x, y \in K$. Let $\pi_0(dxdy)$ be the optimal plan corresponding to μ, ν and c. Then, by definition,

$$c(\mu,\nu) = \int_X \int_Y c(x,y)\pi_0(dxdy) \ \ ; \ \ c_m(\mu,\nu) \le \int_X \int_Y c_m(x,y)\pi_0(dxdy)$$

so

$$\phi^m(\mu,\nu) \le \int_X \int_Y (c_m(x,y) - c(x,y))\pi_0(dxdy) \le \sup|h^{''}|C(K)^2 m^{-2/d} \ ,$$

since π_0 is a probability measure. $\square$

If $h(s) = 2^{\sigma-1}s^\sigma$ (hence $c(x,y) = |x-y|^\sigma$) then the condition of Lemma 5.1 holds if $\sigma \ge 2$. Note that if $\mu = \nu$ then $c(\mu,\mu) = 0$ so $\phi^m(\mu,\mu) = \inf_{Z_m \in Z} c^{Z_m}(\mu,\mu)$. In that particular case we can improve the result of Lemma 5.1 as follows:

Proposition 5.2. *If $c(x,y) = |x-y|^\sigma$, $\sigma \ge 1$, $X = Y = Z = \mathbb{R}^d$ and $\nu = \mu = f(x)dx$*

$$\lim_{m\to\infty} m^{\sigma/d}\phi^m(\mu,\mu) = C_{d,\sigma}\left(\int f^{d/(d+\sigma)}dx\right)^{(d+\sigma)/d} \tag{38}$$

where $C_{d,\sigma}$ is some universal constant.

Proof. From (15), $\Xi^{\mu}_{\mu,Z_m}(\vec{p}) = \Xi^{Z_m}_{\mu}(\vec{p}) + \Xi^{Z_m}_{\mu}(-\vec{p})$ is an even function. Hence its maximizer must be $\vec{p} = 0$. By Theorem 2.1

$$\Xi^{\mu,Z_m}_{\mu}(0) = c^{Z_m}(\mu,\mu) \ .$$

Using (13, 14) with $c^{(1)}(x,y) = c^{(2)}(y,x) = 2^{\sigma-1}|x-y|^\sigma$ we get

$$\Xi^{\mu,Z_m}_{\mu}(0) = 2^\sigma \int_{\mathbb{R}^d} \min_{z\in z_m} |x-z|^\sigma \mu(dx) \ .$$

We then obtain (38) from Zador's Theorem [10, 25, [10]. $\square$

Note that Proposition 5.2 does not contradict Lemma 5.1. In fact $\sigma \ge 2$ it is compatible with the Lemma, and (37) holds with $C(\mu,\mu) = 0$ if $\sigma > 2$. If $\sigma \in [1,2)$, however, then the condition of the Lemma is not satisfied (as $h^{''}$ is not bounded near 0), and the Proposition is a genuine extension of the Lemma, in the particular case $\mu = \nu$.

We can obtain a somewhat sharper result for *any* pair μ, ν in the case $\sigma = 2$, which is presented below.

Let $X = Y = Z = \mathbb{R}^d$, $c(x,y) = |x-y|^2$, μ, ν are Borel probability measures which admits a finite second moment. Assume μ is asbolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$. In that case, Brenier Polar factorization Theorem [3] implies the existence of a unique solution to the quadratic Monge problem, i.e a Borel mapping T such that $T_\#\mu = \nu$. Let $\lambda = f(x)dx$ be the McCann interpolation between μ and ν, that is, $\lambda = (I/2 + T/2)_\#\mu$. We know that λ is absolutely continuous with respect to Lebesgue as well.

Theorem 5.1. *Under the above assumptions,*

$$\limsup_{m\to\infty} m^{2/d}\phi^m(\mu,\nu) \le 4C_{d,2}\left(\int f^{d/(d+2)}dx\right)^{(d+2)/d} \ .$$

Proof. Let S be the optimal Monge mapping transporting λ to ν, i.e. $S_\#\lambda = \nu$ is a solution of Monge problem

$$\int_{\mathbb{R}^d} |S(x)-x|^2\lambda(dx) = \min_{Q;Q_\#\lambda=\nu} \int_{\mathbb{R}^d} |Q(x)-x|^2\lambda(dx) \ .$$

Note that if $y = (T(x)+x)/2$ then $S(y) = T(x)$. Then, since $\lambda = (I/2 + T/2)_\#\mu$,

$$\int_{\mathbb{R}^d} |S(y)-y|^2\, \lambda(dy) = \int_{\mathbb{R}^d} \left|\frac{T(x)-x}{2}\right|^2 \mu(dx) \equiv c(\mu,\nu)/4 \ .$$

Also, if $y = (T(x)+x)/2$ then $2y - S(y) = x$. If follows that $2I - S$ is the optimal Monge mapping transporting λ to μ, that is,

$$\int_{\mathbb{R}^d} |S(x)-2x|^2\lambda(dx) = \int_{\mathbb{R}^d} |S(x)-x|^2\lambda(dx) = \min_{Q;Q_\#\lambda=\mu} \int_{\mathbb{R}^d} |Q(x)-x|^2\lambda(dx)$$

so

$$c(\mu,\nu) = 2\int_{R^d} |S(x)-x|^2\lambda(dx) + 2\int_{R^d} |S(x)-2x|^2\lambda(dx) = 4\int_{R^d} |S(x)-x|^2\lambda(dx) \ . \tag{39}$$

Given $z \in Z_m$, let

$$V_z := \left\{ x \in \mathbb{R}^d; \quad |x-z| \le |x-z^{'}| \quad \forall z^{'} \in Z_m \right\} \ . \tag{40}$$

Since $\cup_{z\in Z_m} V_z = \mathbb{R}^d$ and $\lambda(V_z \cap V_{z^{'}}) = 0$ for $z \neq z^{'}$ then (39) implies

$$c(\mu,\nu) = 4\sum_{z\in Z_m} \int_{V_z} |S(x)-x|^2\lambda(dx) \ . \tag{41}$$

Let $\nu_z := S_\#\lambda\lfloor V_z$, $\mu_z := (2I-S)_\#\lambda\lfloor V_z$. Form Lemma 2.2

$$c^{Z_m}(\mu,\nu) \le 2\left(\sum_{z\in Z_m} \int |x-z|^2\mu_z(dx) + \sum_{z\in Z_m} \int |x-z|^2\nu_z(dx) \right)$$

$$= 2\sum_{z\in Z_m} \int_{V_z} \left\{ |S(x)-z|^2 + |2x-S(x)-z|^2 \right\} \lambda(dx) \tag{42}$$

By the identity

$$4|z-x|^2 = 2\left\{ |S(x)-z|^2 + |2x-S(x)-z|^2 \right\} - 4|S(x)-x|^2 \ .$$

This, together with (41, 42) and (10) implies

$$\phi^m(\mu,\nu) \le 4\sum_{z\in Z_m} \int_{V_z} |x-z|^2\lambda(dx) \ . \tag{43}$$

By (40), $\sum_{z\in Z_m}\int_{V_z} |x-z|^2\lambda(dx) = \int_{\mathbb{R}^d} \min_{z\in Z_m} |x-z|^2\lambda(dx) := \phi(\lambda, Z_m)$. Since (43) is valid for any Z_m we get the result from Zador's Theorem [10, 25, [10]. $\square$

6 Description of the Algorithm

We now spell out the proposed algorithm for approximating of the optimal plan $c(\mu,\nu)$. We assume that c is given by (5). We fix a large numbers n_1, n_2 (not necessarily equal) which characterizes the fine sampling, and much smaller m characterizing the partition order. Then we choose an appropriate sampling: In X we set $\mu_{n_1} := \sum_{i=1}^{n_1} s_i \delta_{x_i}$ for μ and on Y we set $\nu_{n_2} := \sum_{i=1}^{n_2} \tau_i \delta_{y_i}$ for ν.

At the first stage we choose $Z^{(0)} := \{z_1^0, \dots z_m^0\} \in Z^m$, and define

$$\Xi_0(\vec{p}) := \sum_{i=1}^{n_1} s_i \min_{1\le j\le m}[c^{(1)}(x_i, z_j^0) + p_j] + \sum_{i=1}^{n_2} \tau_i \min_{1\le j\le m}[c^{(2)}(z_j^0, y_i) - p_j]$$

Next we choose a favorite method to maximize Ξ_0 on $\mathbb{R}^m$. It is helpful to observe that Ξ_0 is differentiable a.e. on $\mathbb{R}^m$. Indeed, let

$$A_j^0(\vec{p}) := \{i \in (1, \dots n_1),\ c^{(1)}(x_z, z_j^0) + p_j = \min_{1\le k\le m}[c^{(1)}(x_z, z_k^0) + p_k]\}$$

$$B_j^0(\vec{p}) := \{i \in (1, \dots n_2),\ c^{(2)}(z_j^0, y_z) + p_j = \min_{1\le k\le m}[c^{(2)}(z_k^0, y_z) + p_k]\}\ .$$

Then

$$\frac{\partial \Xi_0}{\partial p_j} = \sum_{i\in A_j(\vec{p})} s_z - \sum_{i\in B_j(-\vec{p})} \tau_z$$

provided $A_j(\vec{p}) \cap A_k(\vec{p}) = \emptyset$ and $B_j(-\vec{p}) \cap B_k(-\vec{p}) = \emptyset$ for any $k \neq j$.

Let $\vec{p}_0$ be the maximizer of Ξ_0 on $\mathbb{R}^m$, $A_j^0 := A_j^0(\vec{p}_0)$, $B_j^0 := B_j^0(-\vec{p}_0)$.

At the l step we are given $Z^{(l)} := \{z_1^l, \dots z_m^l\} \in Z^m$, $\vec{p}_l$ the maximizer of

$$\Xi_l(\vec{p}) := \sum_{i=1}^{n_1} s_z \min_{1\le j\le m}[c^{(1)}(x_z, z_j^l) + p_j] + \sum_{i=1}^{n_2} \tau_z \min_{1\le j\le m}[c^{(2)}(z_j^l, y_z) - p_j]$$

and the corresponding $A_j^l := A_j^l(\vec{p}_l)$, $B_j^l := B_j^l(-\vec{p}_l)$ where

$$A_j^l(\vec{p}) := \{i \in (1, \dots n_1),\ c^{(1)}(x_z, z_j^l) + p_j = \min_{1\le k\le m}[c^{(1)}(x_z, z_k^l) + p_k]\}$$

$$B_j^l(\vec{p}) := \{i \in (1, \dots n_2),\ c^{(2)}(z_j^l, y_z) + p_j = \min_{1\le k\le m}[c^{(2)}(z_k^l, y_z) + p_k]\}\ .$$

We define z_j^{l+1} as the minimizer of

$$\zeta \mapsto \sum_{i\in A_j^l} s_z c^{(1)}(x_z, z) + \sum_{i\in B_j^l} \tau_z c^{(2)}(z, y_z) \tag{44}$$

and set $Z^{(l+1)} := \{z_1^{l+1}, \dots z_m^{l+1}\} \in Z^m$. Now

$$\Xi_{l+1}(\vec{p}) := \sum_{i=1}^{n_1} s_z \min_{1\le j\le m}[c^{(1)}(x_z, z_j^{l+1}) + p_j] + \sum_{i=1}^{n_2} \tau_z \min_{1\le j\le m}[c^{(2)}(z_j^{l+1}, y_z) - p_j]\ .$$

From these we evaluate the maximizer $\vec{p}_{l+1}$ the maximizer of Ξ_{l+1} and the sets A_j^{l+1}, B_j^{l+1}.

Remark 6.1. *The maximizer* $\vec{p}_l$ *at the* l *stage can be used as an initial guess for calculating the maximizer* $\vec{p}_{l+1}$ *at the next stage. This can save a lot of iterations where the stages where changes of the centers* $Z^{(l)} \to Z^{(l+1)}$ *is small.*

Using Proposition 5.1 and Corollary 5.1 we obtain a monotone non increasing sequence

$$\Xi_0(\vec{p}_0) \geq \ldots \geq \Xi_l(\vec{p}_l) \geq \Xi_{l+1}(\vec{p}_{l+1}) \ldots \geq c(\mu, \nu) \ .$$

The iterations stop when this sequence saturate, according to a pre-determined criterion.

6.1 Application for quadratic cost

As a demonstration, let us consider the special (but interesting) case of quadratic cost function $c(x, y) = |x - y|^2$ on Euclidean space $X = Y$. We observe the trivial inequality $|x - y|^2 = 2\min_{z \in X}[|x - z|^2 + |y - z|^2]$. Hence we may approximate $|x - y|^2$ by

$$c^{Z_m}(x, y) := 2 \min_{z \in Z_m}[|x - z|^2 + |y - z|^2] \geq |x - y|^2 \quad . \tag{45}$$

So, we use $c^{(1)}(x, z) := 2|x - z_z|^2, c^{(2)}(z_z, y) := 2|y - z_z|^2$.

The updating (44) takes now a simpler form due to Remark 5.1. Indeed, z_j^{l+1} is nothing but the center of mass

$$z_j^{l+1} = \frac{\sum_{i \in A_j^l} s_z x_z + \sum_{i \in B_j^l} \tau_z y_z}{\sum_{i \in A_j^l} s_z + \sum_{i \in B_j^l} \tau_z} \ .$$

7 Some experiments with quadratic cost on the plane

In this section we demonstrate the algorithm for quadratic cost. The pair (μ, X) is always considered to be uniform Lebesgue measure on the unit square $B := \{(x_1, X_2);\ 0 \leq x_1, x_2) \leq 1\}$. It is sampled by an empiric measure of regular grid composed on 400 points $x_1^{(i)} = i/20, x_2^{(j)} = j/20$, $\mu(\{i/20, j/20)\}) = 1/400,\ \ 1 \leq i, j \leq 20$. The image space (Y, ν) is, again, a probability measure on the plane which depends on the particular experiment. The number of centers $m = 10$ and their initial choice is arbitrary within the unit square.

In the first experiments we used a given mapping $T := (T_1, T_2) : B \to \mathbb{R}^2$, and defined (Y, ν) according to $Y = T(B)$, $\nu = T_{\#}\mu$. In that case the naturel sampling is just $(y_1^{(i)}, y_2^{(j)}) = (T_1(i/20), T_2(j/20))$, and $\nu(\{(y_z^{(1)}, y_j^{(2)})\}) = 1/400$.

In all these experiment we used $T_k = \partial\Phi/\partial x_k,\ k = 1, 2$, where $\Phi(x_1, x_2) = 0.5(x_1^2 + x_2^2) + \lambda(\cos(x_1 + 2x_2) - \sin(x_1 - x_2))$. Figs.1-2 shows the saturated result for different values of λ.

Fig 3-**??** show pair of partitions on the X square. The right square is the image under $(\nabla\Phi)^{-1}$ of the partition in the left square. Note that for small values of λ the two partitions looks identical. This is, in fact, what we expect as long as Φ is a convex function. Indeed, the celebrated Brenier's theorem of polar factorization [3] implies just this! For larger values of λ, Φ is not convex and we see clearly the difference between these two partitions.

In the second class of experiments we used different domains for Y (e.g. T shaped, I shaped and A shaped) which are not induced by a mapping. Fig. **??**-**??** display the induced partitions after saturation for different initial choices of the centers z_z. It demonstrates that the saturated partition may depend on the initial choice of the centers.

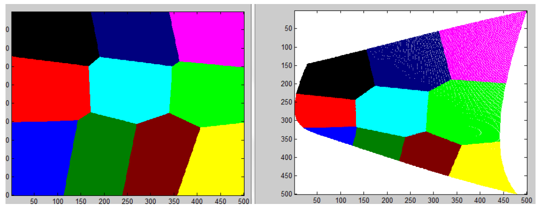

Figure 1: Partition for Φ, $\lambda = 0.2$

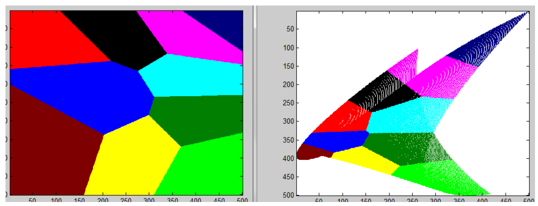

Figure 2: Partition for Φ, $\lambda = 1.2$

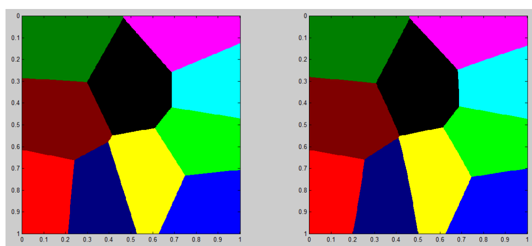

Figure 3: Comparison of partitions $\lambda = 0.05$

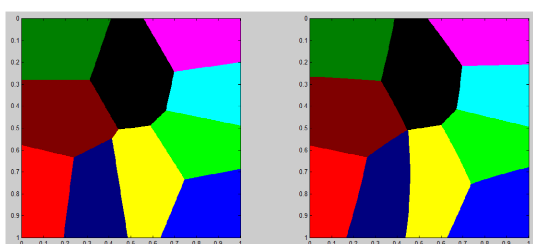

Figure 4: Comparison of partitions $\lambda = 0.1$

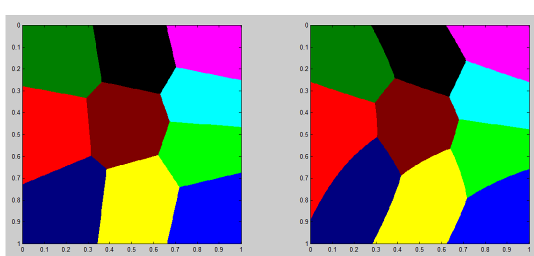

Figure 5: Comparison of partitions $\lambda = 0.2$

8 Comparison with other semi discrete algorithms

Applications of semi-discrete methods for numerical algorithms where introduced in paper by Mérigot [14], followed by a paper of Lévy [4]. Here we indicate the similar and different aspects of our proposed algorithm, compared to [14, 4].

The starting point of Mérigot-Lévy algorithm for quadratic cost involves a discretization ν_m of the target measure ν. For $\nu_m = \sum_1^m r_i \delta_{y_i}$, the optimal plan for transporting μ is obtained by maximizing

$$\mathbb{R}^m \ni \vec{p} \mapsto \int \min_{1\leq i\leq m} [|x-y_i|^2 + p_i]\mu(dx) - \sum_i^m r_i p_i \ . \tag{46}$$

This is equivalent to the function we defined (for the special case of quadratic cost) as $\Xi_\mu^{Z_m}(\vec{p}) - \vec{p}\cdot\vec{r}$, whose maximum over $\mathbb{R}^m$ is $(-\Xi_\mu^{Z_m})^*(-\vec{r})$ as defined in (16). The optimal partition induced by maximizing (46) is refined by taking finer and finer discretization of ν with increasing number of points m. The multi-grid method is, essentially, using the data of the maximizer $\vec{p}$ corresponding to ν_m as an initialization for the $m+1$ level maximization corresponding to (46).

In the present paper we take a different approcah, namely the semi-discretization of the *cost function* $c = c(x, y)$ via (7). It is, in fact, equivalent to a two sided discretization analogus to (46) (in the quadratic case), as we can observe from (19). However, by carrying the duality method one step forward we could reduce the optimization problem to a single one over $\mathbb{R}^m$ via Theorem 2.1.

References

1. L. Ambrosio: *Lecture notes on optimal transport problems Mathematical Aspects of Evolving Interfaces*, Lecture Notes in Math., Funchal, 2000, vol. 1812, Springer-Verlag, Berlin (2003), pp. 1-52

2. Bauschke, H.H and Combettes, P.L: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, (2011).

3. Y. Brenier: *Polar factorization and monotone rearrangement of vector valued functions*, Arch. Rational Mech

4. Bruno Lévy: *A numerical algorithm for L_2 semi-discrete optimal transport in 3D*, arXiv:1409.1279v1 [math.AP] 3 Sep 2014

5. Bogachev, V.I *Measure Theory*, Springer 2007 &Anal., **122**, (1993), 323-351.

6. Caffarelli, L; González, M and Nguyen, T.: *A perturbation argument for a Monge-Ampére type equation arising in optimal transportation*, Arch. Ration. Mech. Anal. 212 (2014), no. 2, 359-414

7. Chiappori, Pierre-André, Robert J. McCann, and Lars P. Nesheim. "Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness." Economic Theory 42.2 (2010): 317-354

8. Frank, A: On Kuhńs Hungarian method-a tribute from Hungary, Naval Research Logistics, 52 (1) (2005), 2-6

9. Glimm, T., and V. Oliker. *Optical design of single reflector systems and the Monge-Kantorovich mass transfer problem*, Journal of Mathematical Sciences 117.3 (2003): 4096-4108

10. Graf, S and Luschgy, H.: *Foundations of Quantization for Probability Distributions*, Lect. Note Math. 1730, Springer, (2000)

11. Kuhn, H.W.: *The Hungarian method for the assignment problem*, Naval Research Logistics Quarterly, 2 (1,2) (1955), 83-97

12. Kuhn, H.W: *Statement for Naval Research Logistics*, Naval Research Logistics, 52 (1) (2005), p. 6

13. Kantorovich, L: *On the translocation of masses*, C.R (Doclady) Acad. Sci. URSS (N.S), 37, (1942), 199-201

14. Mérigot, Q.: *A multiscale approach to optimal transport*, Computer Graphics Forum, Wiley-Blackwell, 2011, 30 (5), pp.1584-1592.

15. Monge,G: *Mémoire sur la théorie des déblais et des remblais*, In Histoire de lÁcadémie Royale des Sciences de Paris, 666-704, 1781

16. Munkres, J:: *Algorithms for the assignment and transportation problems*, Jo Soc. IDUST. APPL. HATtt. Vol. 5, No. 1, March, 1957

17. Papadimitriou, C. H and Kenneth S:. *Combinatorial optimization: algorithms and complexity.* Courier Dover Publications, 1998

18. Pentico, D. W: *Assignment problems: a golden anniversary survey*, European J. Oper. Res. 176 (2007), no. 2, 774-793.

19. Rubinstein, J., Wolansky, G.: *Intensity control with a free-form lens*, J. Opt. Soc. Amer. A 24 (2007), no. 2, 463-469

20. Rubinstein, J., Wolansky, G.: *A weighted least action principle for dispersive waves*, Ann. Physics 316 (2005), no. 2, 271-284

21. Rubner, Y., Tomasi, C. and Guibas, L.: *The Earth Mover's Distance as a Metric for Image Retrieval*, International Journal of Computer Vision, **40**, 2, 99-121 (2000),

22. Villani, C: *Topics in Optimal Transportation*, A.M.S Vol 58, 2003

23. Votaw, D.F and Orden, A: *The personnel assignment problem*, Symposium on Linear Inequalities and Programmng, SCOOP 10, US Air Force, 1952, 155-163

24. Wolansky, G.: *On Semi-discrete Monge Kantorovich and Generalized Partitions*, to appear in JOTA

25. P.L. Zador, *Asymptotic quantization error of continuous signals and the quantization dimension*, IEEE Trans. Inform. Theory 28, Special issue on quantization, A. Gersho & R.M. Grey Eds. (1982) 139-149.

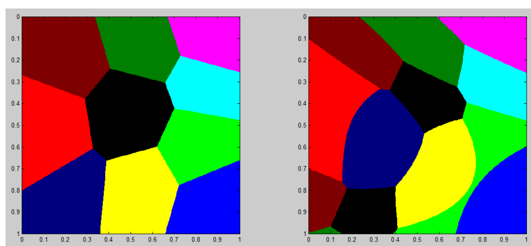

Figure 6: Comparison of partitions $\lambda = 0.5$